Innovent
信达生物制药

STATISTICAL ANALYSIS PLAN

A Randomized, Double-Blind, Multi-Center Phase III Clinical Trial Evaluating the Efficacy and Safety of IBI308 or Placebo in Combination with Oxaliplatin and Capecitabine (XELOX), for First-Line Treatment of Unresectable, Locally Advanced, Recurrent, or Metastatic Gastric or Gastroesophageal Junction Adenocarcinoma (ORIENT-16)

Protocol No.: CIBI308E301

Version: Final Version 1.0

Version Date: 30Jul2021

Statistician: [redacted]

VERSION HISTORY:

Version	Version date	Main amendments
0.1	30May2020	First draft
1.0	30Jul2021	Final version

SIGNATURE AND APPROVAL PAGE

I totally agree with this statistical analysis plan, and confirm that the plan covers all the clinical trial objectives specified in the protocol.

Author:

_____ Date: _____

[redacted]

Approved by:

_____ Date: _____

[redacted]

_____ Date: _____

[redacted]

ABBREVIATIONS

Abbreviation	Full name
ADA	Anti-drug antibody
AE	Adverse event
ATC	Anatomical therapeutic chemical classification system
BMI	Body mass index
BSA	Body surface area
CPS	Combined positive score
CR	Complete response
CRF	Case report form
CTCAE	Common terminology criteria for adverse events
DCR	Disease control rate
DoR	Duration of response
e-CRF	Electronic case report form
ECG	Electrocardiogram
ECOG	Eastern cooperative oncology group
ECOG PS	Eastern cooperative oncology group performance status
EORTC QLQ	European organization for research and treatment of cancer quality of life questionnaire
EOS	End of study
EQ-5D-5L	European quality of life – 5 dimensions – 5 levels
HR	Hazard ratio
iDCC	Independent data coordinating center
iDMC	Independent data monitoring committee
ITT	Intention to treat
IWRS	Interactive web response system
MDD	Minimum detectable difference
MedDRA	Medical dictionary for regulatory activities
mITT	Modified intention to treat
mmHg	Millimetres of mercury
NAb	Neutralizing antibody
ORR	Overall response rate
OS	Overall survival
PD	Progressive disease
PD-L1	Programmed death-ligand 1
PFS	Progression free survival
PPS	Per-Protocol set
PK	Pharmacokinetics
PR	Partial response
PT	Preferred term
RECIST	Response evaluation criteria in solid tumors
SAE	Serious adverse event

Abbreviation	Full name
SAP	Statistical analysis plan
SAS	Statistical analysis system
SD	Stable disease
SOC	System organ class
SS	Safety analysis set
TEAE	Treatment-emergent adverse events
TPS	Tumor proportion score

TABLE OF CONTENTS

VERSION HISTORY: 2
SIGNATURE AND APPROVAL PAGE 3
ABBREVIATIONS 4
TABLE OF CONTENTS 6
1 INTRODUCTION 9
2 SUMMARY OF KEY INFORMATION IN PROTOCOL 9
2.1 Changes in Planned Statistical Analysis from Protocol 9
2.2 Study Objectives and Endpoints 9
2.3 Study Design 10
3 STATISTICAL CONSIDERATION 12
3.1 Statistical Hypothesis 12
3.2 Sample Size Estimation 13
3.3 Randomization and Blinding 13
3.4 Interim Analysis 14
3.5 Multiplicity 14
3.6 Analysis Population 15
4 GENERAL PRINCIPLES FOR STATISTICAL ANALYSIS AND DATA PROCESSING RULES 16
4.1 General Principles 16
4.2 Processing of Missing Data 16
4.3 Visit Window 18
4.4 Scheduled and Unscheduled Visits 19
4.5 Processing of Other Data 19
4.6 General Definition 19
5 ANALYSIS OF STUDY POPULATION 20
5.1 Subject Disposition 20
5.2 Protocol Deviations 20
6 DEMOGRAPHICS AND BASELINE CHARACTERISTICS ANALYSIS 21
6.1 Demographics and Baseline Characteristics 21

6.2 Previous Therapy of Study Disease 23
6.3 Non-Tumor Medical History 24
6.4 Previous and Concomitant Medication 25
6.5 Previous and Concomitant Non-drug Therapy 25
6.6 Subsequent New Anti-Tumor Therapy 26
7 EFFICACY ANALYSIS 27
7.1 Primary Efficacy Analysis 27
7.1.1 Definition of primary efficacy endpoints OS 27
7.1.2 Primary statistical analysis methods 27
7.1.3 Sensitivity analysis 28
7.1.4 Subgroup analysis 28
7.2 Secondary Efficacy Analysis 28
7.2.1 Secondary efficacy analysis 28
7.2.2 Progression-free survival (PFS) 29
7.2.3 Objective response rate (ORR) 30
7.2.4 Disease control rate (DCR) 31
7.2.5 Duration of response (DoR) 31
8 SAFETY ANALYSIS 31
8.1 Investigational Drug Exposure 31
8.2 Adverse Events 34
8.2.1 Definition 34
8.2.2 Planned analysis 34
8.3 Laboratory Tests 37
8.4 12-Lead ECG 38
8.5 Vital Signs, Physical Examination, and Other Safety Examinations 39
8.5.1 Vital signs 39
8.5.2 Physical examination 40
8.5.3 ECOG PS score 40
8.5.4 Other safety examinations 40
8.6 Immunogenicity Analysis 40

9 EXPLORATORY STATISTICAL ANALYSIS 41
9.1 Analysis of Predictive Biomarkers in Tumor Tissue and Peripheral Blood 41
9.2 Quality of Life Analysis 41
9.2.1 European quality of life - 5 dimensions - 5 levels (EQ-5D-5L) 41
9.2.2 European organization for research and treatment of cancer quality of life questionnaire (EORTC QLQ-C30) 42
9.2.3 Quality-of-life questionnaire in patients with gastric cancer (EORTC QLQ-STO22) 43
10 REFERENCES 45

LIST OF TABLES

Table 1. Definitions of study objectives and endpoints 9
Table 2. Definition of analysis sets 15
Table 3. Events and Censoring rules for primary and sensitivity analysis of PFS 29
Table 4. Dose and administration 33
Table 5. Laboratory parameters 38
Table 6. Potential clinically significant abnormalities of ECG 39
Table 7. Potential clinically significant abnormalities of vital signs 39
Table 8. Algorithm of Calculating Raw Score for QLQ-C30 scale 42
Table 9. Algorithm of Calculating Raw Score for QLQ-STO22 scale 44

LIST OF FIGURES

Figure 1. Schematic of CIBI308E301 study design and administration 12

1 INTRODUCTION

This statistical analysis plan (SAP) is prepared according to the study protocol CIBI308E301: " A Randomized, Double-Blind, Multi-Center Phase III Clinical Trial Evaluating the Efficacy and Safety of IBI308 or Placebo in Combination with Oxaliplatin and Capecitabine (XELOX), for First-Line Treatment of Unresectable, Locally Advanced, Recurrent, or Metastatic Gastric or Gastroesophageal Junction Adenocarcinoma (ORIENT-16)" (Version No.: 3.2, Version date: Jun. 11, 2021). This document provides the design and objectives of the protocol, and provides the detailed definitions of the endpoints and the planned statistical analyses performing for the endpoints. The final SAP will be approved and signed before database lock.

This document describes the efficacy and safety analysis plan, excluding pharmacokinetic (PK) analysis. The pharmacokinetic (PK) analysis plan will be presented in a separate file.

2 SUMMARY OF KEY INFORMATION IN PROTOCOL

2.1 Changes in Planned Statistical Analysis from Protocol

There is no change in this SAP from the study protocol CIBI308E301 (Jun. 11, 2021, Version 3.2).

2.2 Study Objectives and Endpoints

Table 1. Definitions of study objectives and endpoints

Study objectives	Study endpoints
Primary objectives	Primary endpoints
• To compare the overall survival (OS) of Sintilimab vs. placebo in combination with chemotherapy, for first-line treatment of unresectable, locally advanced, recurrent, or metastatic gastric or gastroesophageal junction adenocarcinoma • To compare the OS of Sintilimab vs. placebo in combination with chemotherapy, for first-line treatment in PD-L1-positive subjects with unresectable, locally advanced, recurrent, or metastatic gastric or gastroesophageal junction adenocarcinoma	• OS in ITT population • OS in PD-L1-positive subjects in ITT population
Secondary objectives	Secondary endpoints
• To compare the progression-free survival (PFS) of the subjects between the two groups	• Progression-free survival (PFS)

Study objectives	Study endpoints
• To compare the objective response rate (ORR) between the two groups	• Objective response rate (ORR)
• To compare the disease control rate (DCR) between the two groups	• Disease control rate (DCR)
• To compare the duration of response (DoR) between the two groups	• Duration of response (DoR)
• To compare the safety between the two groups	• Incidence, relationship, and severity of all adverse events (AEs), treatment-emergent adverse events (TEAEs), treatment-related adverse events (TRAEs), serious adverse events (SAEs), and immune-related adverse events (irAEs) • Changes in ECG, vital signs, physical examination results, and laboratory results before, during, and after treatment • Incidence of anti-drug antibody (ADA) and neutralizing antibody (NAb) positive results
Exploratory objectives	Exploratory endpoints
• To compare changes in quality of life between the two groups	• EQ 5D-5L scale • EORTC QLQ-C30 scale • EORTC QLQ-STO22 scale
• To evaluate the pharmacokinetic (PK) characteristics of Sintilimab in combination with chemotherapy in patients with unresectable, locally advanced, recurrent, or metastatic gastric or gastroesophageal junction adenocarcinoma	• Including, but not limited to, the trough concentration of Sintilimab in cycles 1/3/11
• To evaluate the correlation between biomarkers in tumor tissue and efficacy, including PD-L1 expression level	• PD-L1 expression in tumor tissues

2.3 Study Design

This is a randomized, double-Blind, multi-center phase III clinical trial evaluating the efficacy and safety of IBI308 or Placebo in combination with oxaliplatin and capecitabine (XELOX), for first-line treatment of unresectable, locally advanced, recurrent, or metastatic gastric or gastroesophageal junction adenocarcinoma.

Subjects will be treated with Sintilimab (weight < 60 kg: 3 mg/kg IV Q3W; weight ≥ 60 kg: 200 mg IV Q3W) or placebo, in combination with XELOX regimen (Oxaliplatin 130 mg/m^2 IV Q3W, and capecitabine 1000 mg/m^2 Bid PO × 14d Q3W), for up to 6 cycles of combination chemotherapy (1 cycle = 3 weeks). Then the subjects will receive Sintilimab or placebo, in combination with capecitabine (1000 mg/m^2 Bid PO × 14d Q3W), for maintenance treatment until PD, intolerable toxicity, initiation of new anti-tumor therapy, withdrawal of ICF, loss to follow-up, death, or any other reason for treatment discontinuation judged by the investigator,

whichever comes first. Treatment with sintilimab or placebo, in combination with Capecitabine, will last for up to 24 months (starting from the first dose). If a drug is discontinued for any reason during the treatment, other drugs are permitted to be continued.

Subjects with unresectable, locally advanced, recurrent, or metastatic gastric or gastroesophageal junction adenocarcinoma will be randomized to sintilimab group or placebo group in a 1:1 ratio, 325 in the sintilimab arm and 325 in the placebo arm, totally 650 subjects. The randomization stratification factors include the Eastern Cooperative Oncology Group Performance Status (ECOG PS) score (0 or 1), hepatic metastasis (yes or no), and PD-L1 expression (CPS < 10 or ≥ 10) (CPS refers to combined positive score, which is the sum of PD-L1 expression in infiltrating lymphocytes in tumor cells and tumor tissues). The primary endpoint of the study is the OS in the ITT population or in the PD-L1 positive subjects (CPS ≥ 5). OS is defined as the time from randomization to death for any cause. Subjects who are still alive at the time of analysis will be censored at the last known date of survival or the analysis cut off date, whichever is the earliest.

In this study, clinical tumor imaging evaluation will be performed according to RECIST v1.1. During the study, tumor imaging evaluation will be performed once every 6 weeks (± 7 days) for 48 weeks, and then once every 12 weeks (± 7 days) until PD, initiation of new anti-tumor therapy, withdrawal of ICF, loss to follow-up, death, or end of study, whichever occurs first.

An interim safety analysis and an interim efficacy analysis will be performed during the study, and the results and reports will be submitted to the independent data monitoring committee (iDMC). The iDMC will review the safety data in the interim safety analysis, and provide an advice to the sponsor on whether the study can be continued, and make recommendation to sponsor whether the trial data can be submitted earlier based on the pre-specified efficacy boundaries. Prior to the interim analysis, iDMC charter will be finalized and approved by the iDMC and the sponsor. The responsibilities of iDMC members and related procedures will be defined in iDMC charter.

After the study treatment is terminated and completed, safety follow-up and survival follow-up will be performed (once every 60 days).

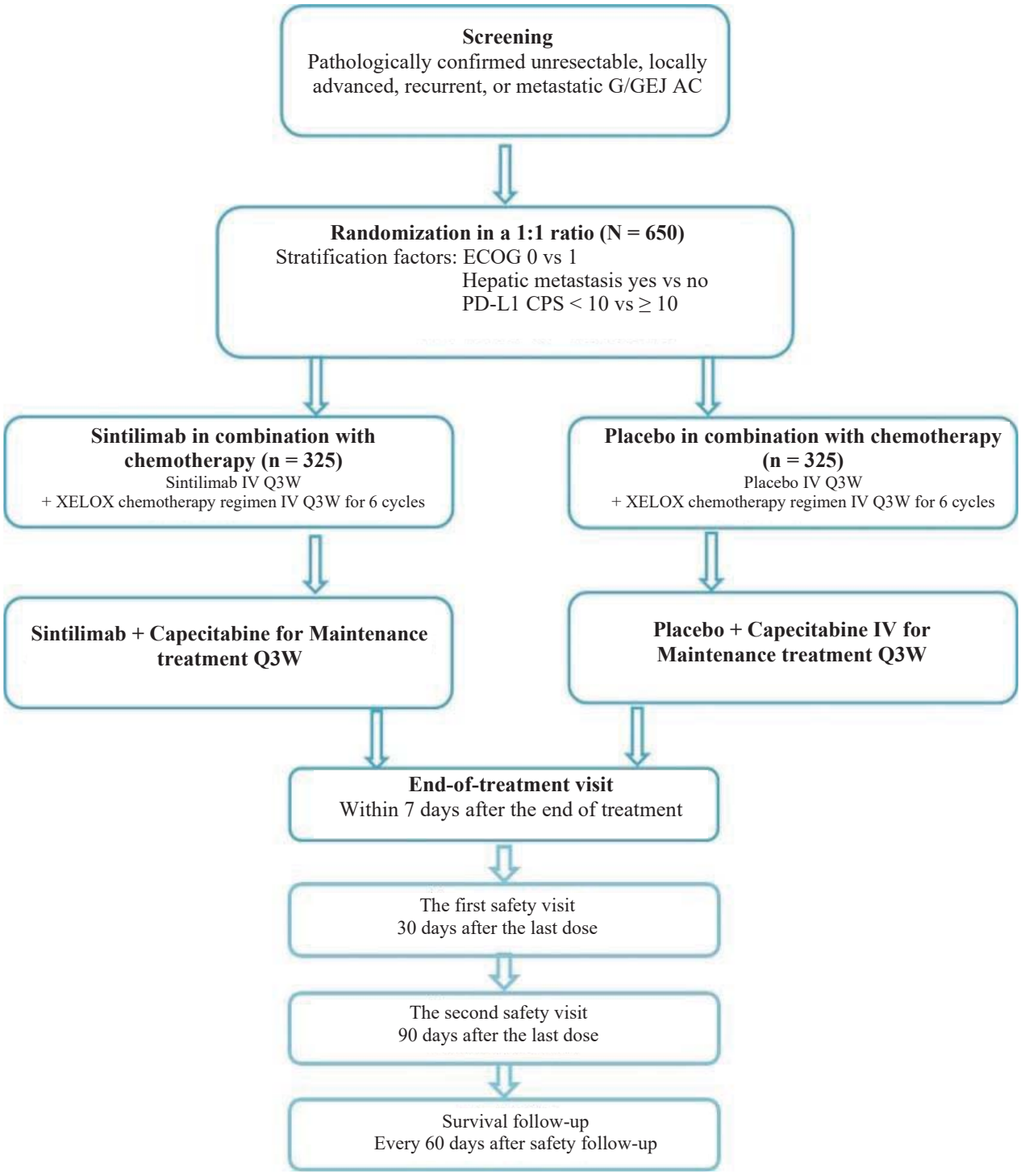

Figure 1. Schematic of CIBI308E301 study design and administration

3 STATISTICAL CONSIDERATION

3.1 Statistical Hypothesis

This is a superiority trial. The primary efficacy endpoints are the OS in ITT population and in the PD-L1-positive population. The superiority hypothesis test is:

Null hypothesis H_0: HR ≥ 1

Alternative hypothesis H_a: HR < 1

In this study, the test will be performed in a fixed order: first, the superiority test will be performed for OS in PD-L1 positive population, and then the test will be performed for OS in ITT population when the OS in PD-L1 positive population reaches statistical significance.

3.2 Sample Size Estimation

For OS in the PD-L1 positive population, assuming that the hazard ratio (HR) of Sintilimab to placebo, in combination with chemotherapy, is 0.7 (median OS is 15.7 and 11 months, respectively), 287 OS events are required to achieve a statistical test power of about 85% at the two-sided significance level of 0.05. For OS in ITT population, assuming that the HR of Sintilimab to placebo, in combination with chemotherapy, is 0.75 (median OS is 14.7 months and 11 months, respectively), 515 OS events are required to achieve a statistical test power of about 90% at the two-sided significance level of 0.05.

Based upon the monthly censoring rate of 0.5%, assuming that the study will take 18 months to enroll 650 subjects, 325 in each group, 515 OS events are estimated to be observed within 46 months in ITT population. Based on the above assumptions, PD-L1 positive population accounts for 56.2% (i.e. 365 subjects) of ITT population, and 287 OS events can be observed at the final analysis in PD-L1 positive population.

3.3 Randomization and Blinding

In this study, the randomization stratification factors are as following:

- ECOG PS score (0 or 1)
- Hepatic metastasis (yes or no)
- PD-L1 expression (CPS < 10 or ≥ 10).

Subjects will be randomized at a 1:1 ratio to the test group and the control group, receiving Sintilimab in combination with chemotherapy, and placebo in combination with chemotherapy, respectively. During randomization, study site will not be regarded as a randomization stratification factor, and all sites will compete for enrollment through the IWRS. A randomization number will be generated by the IWRS and assigned to each subject who has completed all of the screening procedure and meets the inclusion criteria. If a randomized subject

withdraws from the study for any reason, the randomization number will be retained. Randomization statistician will only develop the randomization list used for randomization, and will not be involved in the conduct of the study.

Sintilimab and placebo will be in the same packages to ensure blindness. Except for emergency unblinding outlined in Protocol 8.8.1, the investigator, subjects, and sponsor remain blind to the treatment allocations during the study.

3.4 Interim Analysis

One safety interim analysis and one efficacy interim analysis will be performed in this study.

The safety interim analysis will be performed when about 200 PFS events are observed in the study, to monitor the overall safety of clinical trial subjects. The sponsor will remain blind during the safety interim analysis. Unblinded results of the interim analysis will be generated by the independent Data Coordinating Center (iDCC) and delivered to the iDMC.

Meanwhile, an interim efficacy analysis is planned for this study. The fixed sequence test will be used for hypothesis testing: first, the superiority test will be performed for OS in PD-L1 positive population, and then the test will be performed for OS in ITT population when the OS in PD-L1 positive population reaches statistical significance. The interim efficacy analysis of OS in the two populations will be performed when at least 70% of total OS events have been reached in the PD-L1-positive population and the overall population (i.e., 361 OS events in ITT population, and 201 OS events in PD-L1 positive population), and the test level will follow the Lan-Demets approach to the O'Brien-Fleming boundary. In the interim analysis, the nominal test level is 0.0148 (two-sided), the minimum detectable difference (MDD) of OS in PD-L1 positive population is HR = 0.709, and the minimum detectable difference (MDD) of OS in ITT population is HR = 0.774. At the final analysis, the nominal test level is 0.0455 (two-sided), the minimum detectable difference (MDD) of OS in PD-L1 positive population is HR = 0.790, and the minimum detectable difference (MDD) of OS in ITT population is HR = 0.838. The exact α value of OS analysis will be adjusted according to the number of OS events occurring in real time to approximate the O'Brien-Fleming boundary according to the Lan-DeMets method, so as to ensure the overall detection level of OS $\alpha = 0.05$.

3.5 Multiplicity

In this study, the fixed sequence test will be used for hypothesis test, and the overall test level will be controlled to $\alpha = 0.05$ (two-sided). For the primary endpoint OS, the superiority test will be performed for OS in PD-L1 positive population, and then the test will be performed for OS in ITT population when the OS in PD-L1 positive population reaches statistical significance. The

overall type I error for hypothesis testing on both populations for the efficacy endpoints of OS was strictly controlled through a fixed order testing approach.

The key secondary efficacy endpoints will be tested at a test level $\alpha = 0.05$ (two-sided) only when the primary efficacy endpoints are statistical significant.

3.6 Analysis Population

Table 2. Definition of analysis sets

Analysis Set	Definition/Criteria	Endpoints of analysis
Intention-to-treat (ITT) set	• All randomized subjects. • Treatment group assigned at randomization will be used during the analysis.	• Efficacy endpoints other than ORR, DCR, and DoR • Subject disposition • Protocol deviations • Medical history • Previous and concomitant medications • Analysis of demographics and baseline characteristics
Modified intention to treat (mITT) set	• All randomized subjects who have measurable lesions at baseline. • Treatment group assigned at randomization will be used during the analysis.	• ORR, DCR, and DoR
Per-protocol set (PPS)	• Subjects who have no major protocol deviations that affect the efficacy evaluation and have completed the minimum exposure to study treatment. • A subset of ITT • PPS will be defined in the blind review meeting before the database lock. The sponsor or its designee will be responsible for reviewing the data of major protocol deviations in a blinded manner, and the analysis population will be defined before the database lock.	• Sensitivity analyses of primary efficacy endpoints and key secondary efficacy endpoints
Safety set (SS)	• All randomized subjects who have received at least one dose of investigational drug. The subjects who have not been randomized but received the investigational drug are not included in the SS. The randomized subjects for whom the treatment with the investigational drug is not determined should be included in the SS, and the treatment group assigned at randomization will be used as the actual treatment arm.	• Safety analysis • Immunogenicity analysis • Investigational drug exposure analysis

Analysis Set	Definition/Criteria	Endpoints of analysis
	• The analysis will be performed based on the actual medication received by the subjects. As long as the subject has received IBI308 once, the actual treatment group will be considered IBI308 in combination with chemotherapy group.	

4 GENERAL PRINCIPLES FOR STATISTICAL ANALYSIS AND DATA PROCESSING RULES

4.1 General Principles

All statistical analyses will be performed using SAS v9.4 or above.

Unless otherwise specified, the data collected by CRF in this study will be presented in the form of a data list and summarized in the table by treatment groups. Categorical data will be summarized by the number and percentage of subjects in each category. Unless otherwise specified, the number of subjects in each treatment group in the corresponding study population will be used as the denominator of the percentages. Percentages will be rounded to 1 decimal place.

The quantitative data will be summarized using the number of cases, mean, standard deviation, median, minimum, and maximum. Mean and median values will have 1 more decimal place than the source data, the standard deviation will have 2 more decimal places than the source data, and the maximum and minimum values will be rounded to the same decimal places as the source data do. The hazard ratio will be rounded to 3 decimal places. *p* values will be rounded to 4 decimal places (or > 0.9999, or < 0.0001).

4.2 Processing of Missing Data

Missing data imputation will be only performed in the derived variables of the summary table and will not be presented in the data list. The missing data will be imputed according to the following rules:

- For the missing date of baseline data, the imputation should be performed by the following methods when calculating the duration:
  - If the start date is missing, the month and day will be imputed to be "January" and "01", respectively.
  - If the end date is missing, the missing month will be imputed as "December", and the missing day will be imputed as the last day of the corresponding month.

- If the date is neither the start date nor the end date, the missing day or month will be imputed following the same rules as the missing start date, the month and day will be imputed to be "January" and "01", respectively.
- The date with missing year, month, and day data will not be imputed, and will be treated as a missing value.

- The adverse events with missing time will be checked. The following rules will be used for imputation for the missing date of the adverse events:
  - If the start date is missing, the month and day will be imputed to be "January" and "01", respectively.

    However, if the missing date is the same as the year and / or month of the first dose, the adverse event may occur during the drug administration. If it does not conflict with the end date, the date of the first dose will be used as the start date of the adverse event. If the end date is earlier than the date of the first dose, the month and day of the start date of the adverse event will be imputed as "January" and "01" based on above methods.
  - If the start date is completely missing (i.e. year, month, and day are all missing), then the date of the first dose will be used as the start date if it does not conflict with the end date. If the end date is earlier than the date of the first dose, then the end date will be used to impute the start date of the adverse event.
  - For the end date, the missing month will be recorded as "December", and the missing day will be recorded as the last day of the corresponding month. If the date obtained by this algorithm is after the date of death, data cutoff, or end of study, the date of death, data cutoff, or end of study, whichever occurs first, will be used for imputation.
  - For an end date with missing year, month, and day data, the corresponding adverse event should be considered to be “ongoing”.
  - The missing dates recorded in the eCRF record will be provided in the listing.
- The CTCAE Grade of adverse events should not be missing. If it is missing, it will be reviewed item-by-item. If there still are missing data, this AE will be considered as CTCAE Grade 3 and above AE when summarizing these AEs. When summarizing by Grade, the "missing row/column will be presented.
- If the relationship with the investigational drug is missing, then the AE will be considered as related to the investigational drug, and no data imputation is required.

- If the first dose date is missing, all the AEs that occurred on or after the date of randomization will be considered as TEAEs.
- The missing dates of previous and concomitant medications and non-drug treatments will be imputed according to the following rules:
  - The missing start and end dates of previous and concomitant medications and non-drug treatments will not be imputed. A missing case that can not be confirmed to be previous or concomitant medication will be considered to be concomitant.
  - If duration is required for a previous medication / concomitant medication / non-drug therapy and the end time of the therapy is missing, the date of death, data cutoff, or end of study, whichever occurs first, will be used for imputation.
- The missing death date will be imputed according to the following rules:
  - The missing day of a death date will be imputed with the first day of that month; unless the known last date of subject survival is in the same month as the death date, in this case, it will be imputed as the known last survival date + 1 day;
  - The missing month and day of a death date will be respectively imputed with "January" and "1" of that year; unless the known last date of subject survival is in the same year as the death date, in this case, they will be imputed as the known last survival date + 1 day;
  - The death date with missing day, month, and year data will be imputed as the known last survival date + 1 day;
- The missing date of new anti-tumor treatment will be imputed as follows:
  - The missing day will be imputed as the last day of that month; unless being in the same year and month as a known study end date collected on the EOS page, then it will be imputed as the end date of the study.
  - The missing month and day will be imputed as the last month and day of that year; unless being in the same year as a known study end date collected on the EOS page, then they will be imputed as the end date of the study.
- Unless otherwise specified, the missing values of other data points will not be imputed, and only the observed values will be used for data analysis and listing.

4.3 Visit Window

This study will be analyzed according to the visit names collected in the CRF.

4.4 Scheduled and Unscheduled Visits

The summary analysis by time point only includes scheduled visits, and unscheduled visits will be included in the summary table of baseline characteristic variable derivation and any post-baseline status, but not in the summary analysis by time point. Unscheduled visits will be presented in the listing.

4.5 Processing of Other Data

If the reported values of laboratory parameters cannot be used in the statistical summary table (for example, the values of numerical parameters are reported as strings), the coded values will be properly determined and applied in statistical analysis. Generally, the upper and lower limits of the normal range, such as "< 10" or "≤ 5", will be treated as "10" or "5" respectively, and "< 100" will be treated as "100" for analysis and summary. Nevertheless, the actual values reported in the database will be presented in the listing.

4.6 General Definition

Age

The calculation of age will be based on the date of signing of the informed consent form from the date of birth. The calculation formula is:

Integer [(date of signing ICF – date of birth + 1) / 365.25].

Baseline

Unless otherwise specified, the last non-missing value prior to the first dose of investigational drug will be used as baseline value. For a subject who is randomized but takes no investigational drug, the baseline value is defined as the last non-missing value before randomization.

Change from baseline

The change from baseline will be calculated for each subject at given time points by subtracting the baseline value from the value at that specific time point.

Relative Day of study

Relative day of study will be calculated as the number of days relative to the date of first dose:

If the date is missing, the relative day of study are missing;

If the date is earlier than date of first dose, the relative day of study = date – date of first dose;

If the date is no earlier than date of first dose, the relative day of study = date – date of first dose + 1.

5 ANALYSIS OF STUDY POPULATION

5.1 Subject Disposition

The summary of subject disposition will include the following information:

- The number of subjects screened.
- Summary of subject disposition in each study site.
- The number and percentage of subjects randomized, subjects randomized and receiving the investigational drug, subjects randomized but not receiving the investigational drug, subjects still receiving the investigational drug, subjects who end of treatment and the primary reasons, subjects who end of the study and the primary reasons by treatment groups. Subject survival assessments: survival, death, or unknown. The subject survival assessments are based on the information on the survival follow up page. During the interim analysis, the survival status of subjects who are still treated in the trial is considered as survival.
- Follow-up time of OS (calculated by the inverse of the censoring rules in the primary OS analysis).
- List of all subjects who end of the treatment or end of study according to the CRF record.
- The number and percentage of subjects in each analysis dataset (ITT, mITT, PPS, and SS) by the treatment groups. List of the reasons for exclusion of subjects from each analysis dataset.
- List of discrepancies between actual treatment and the assigned treatment during randomization for subjects.
- A summary table for the analysis set of PD-L1 positive population (CPS $\geq$ 5) will be provided.

5.2 Protocol Deviations

Major protocol deviations (including protocol deviations related to inclusion / exclusion criteria, the conduct of the study, subject management or subject assessments) will be summarized and a listing of all protocol deviations will be provided too.

The protocol deviations will be recorded by the study team according to deviation management plan in the protocol during the study. All data will be reviewed before database freezen to ensure that all major protocol deviations and deviations that may lead to the exclusion of subjects from

the analysis datasets can be collected and classified in the protocol deviation dataset. This dataset will be the basis for summarizing and listing protocol deviation information.

The subjects meeting the following criteria will be excluded from PPS in this trial:

- Protocol deviations:
  - Subjects who are found not meeting the primary inclusion criteria or meeting the primary exclusion criteria after randomization
    - Primary inclusion criteria include: Inclusion criteria 1, 3, 4.
    - Primary exclusion criteria include: Exclusion criteria 2, 3, 4, 7–12, 16–21, 23, 24, 37, 38.

      Note: The subjects are excluded from the PPS after confirmed by medical experts that the violation of the inclusion and exclusion criteria affects the efficacy evaluation.
  - Subjects who take the prohibited drugs specified in the trial protocol that are evaluated to affect the efficacy by medical experts (see Section 5.4.1 in the protocol for prohibited drugs).
  - Subjects whose take actual treatment inconsistent with the assigned treatment. For subjects taking the wrong drug kit number, the actual drug will be identified after unblinding.
- Subjects who don't complete pre-specified minimal exposure. The definition of minimal exposure: at least one injection of IBI308 or placebo, or at least one injection of Oxaliplatin, or the actual number of days of taking Capecitabine is no less than 14 days.

6 DEMOGRAPHICS AND BASELINE CHARACTERISTICS ANALYSIS

6.1 Demographics and Baseline Characteristics

The following demographics and baseline characteristics collected prior to the first dose of investigational drug will be descriptively summarized by treatment groups:

- Gender
- Age = (date of ICF – date of birth + 1)/365.25, rounded to the smallest integer.
- Age group (< 65 years, ≥ 65 years)
- Race

- Ethnicity
- Height (cm)
- Weight (kg)
- Weight group (< 60 kg, ≥ 60 kg)
- Body mass index (BMI) (kg/m^2) = weight (kg)/(height (m))^2
- Body surface area (BSA) (m^2)
- Random stratification factors: actual stratification factors and stratification factors captured in randomization system.
- PD-L1 expression level: CPS < 5 or CPS ≥ 5, CPS < 1 or CPS ≥ 1, TPS < 10% or TPS ≥ 10%, TPS < 5% or TPS ≥ 5%, TPS < 1% or TPS ≥ 1%
- Family history: with or without family history of tumor
- History of alcohol: alcohol use status (never, current drinker and abstained from drinking), duration of alcohol use (month) and average daily volume of alcohol use (g) for subjects who are currently drinking and who have abstained from drinking
- History of smoking: smoking status (never, current smoker and abstained from smoking), duration of smoking (month) and average number of cigarettes smoked per day for subjects who are currently smoking and who have abstained from smoking
- History of helicobacter pylori (HP) infection
    - HP infection test results: positive, negative, or unknown
    - Test methods: ^{13}C-urea breath test, ^{14}C-urea breath test, rapid urease test, pathological section staining, HP antigen serologic testing
    - Whether the HP infection is treated: Yes, No
- Baseline tumor characteristics
    - Primary diagnosis of tumor:
        - Pathological diagnosis: adenocarcinoma, hepatoid adenocarcinoma, mucinous adenocarcinoma, signet ring cell carcinoma, others
        - Sites of disease: gastric, gastroesophageal junction, others
        - Pathological diagnosis methods: surgery, biopsy

- ICH test results: 0, 1+, 2+, 3+, not done
- FISH test results: positive, negative, not done
- T staging at screening: Tx, T0, Tis, T1, T1a, T1b, T1c, T2a, T2b, T3, T4
- N staging at screening: Nx, N0, N1, N2, N3, N3a, N3b
- M staging at screening: Mx, M0, M1
- Clinical staging at screening: III, IVA, IVB
- G staging at screening: Gx, G1, G2, G3
- Types of disease: locally progressive, metastatic, others
- Whether the tumor has metastasized
- Whether has metastasized to distant lymph nodes, liver, lung, brain, bone, peritoneum, adrenal gland, colon and rectum, kidney, ureter, ovary or other organs

o Tumor assessment of baseline target lesions and non-target lesions

The above demographics and baseline characteristics will also be summarized in the PD-L1 positive population (CPS ≥ 5).

The demographics and other baseline characteristics will also be presented in the listing.

6.2 Previous Therapy of Study Disease

Previous therapy of study disease includes the history of surgery, chemotherapy and radiotherapy.

- Surgery history
  - o The number and percentage of subjects with at least one surgery related to the study disease will be summarized
  - o Previous radical surgery: Yes, No; reviewed, identified and flagged by medical experts
  - o The number and percentage of subjects receiving surgery will be summarized by surgical sites
  - o The verbatim surgery name will be coded according to MedDRA dictionary (version 23.0 or above), and the number and percentage of subjects will be summarized by SOC and PT. For the same SOC or PT, multiple surgeries of the same subject will be counted once only.

- Chemotherapy history
  - The number and percentage of subjects with at least one chemotherapy will be summarized
  - Reasons for receiving chemotherapy: adjuvant chemotherapy, neoadjuvant chemotherapy, others
  - Reasons for treatment discontinuation: increased toxicity, have completed course of treatment, disease progression, others
  - Treatment response: complete response, partial response, stable disease, progressive disease, not applicable, not evaluable
  - The verbatim term of chemotherapy drugs will be coded according to the WHO-Drug dictionary (version: WHODrug Global 2020 Mar or above).The number and percentage of subjects will be summarized by ATC level 2, generic name and treatment group. For the same ATC classification and generic drug name, multiple drugs will be counted once only for a given subject.
- Radiotherapy
  - The number and percentage of subjects with at least one radiotherapy
  - Radiotherapy sites: brain, lung, adrenal gland, lymph node, pleura, liver, bone, skin, others
  - Radiotherapy method: X-ray, γ ray, others
  - Total irradiation dose (Gy): the number of cases, mean, standard deviation, median, minimum, and maximum.
  - Purpose of radiotherapy: radical, palliative, symptomatic, others
  - Treatment response: complete response, partial response, stable disease, progressive disease, not applicable, not evaluable

Previous therapy of study disease will be provided in the listing.

6.3 Non-Tumor Medical History

The non-tumor medical history will be coded according to MedDRA dictionary (version 23.0 or above), and will be summarized by SOC, PT (if applicable) and treatment group. For the same SOC and PT, multiple histories of the same subject will be counted once only.

The non-tumor medical history will be provided in the listing.

6.4 Previous and Concomitant Medication

Previous medication

A previous medication is defined as a medication that has been taken before the first dose of the investigational drug and has been discontinued before the first dose of the investigational drug.

Concomitant medication started before the first dose of the investigational drug

It is defined as a medication that has been taken before the first dose of the investigational drug and is continued after the first dose of the investigational drug. A medication with unknown end date or dose continuation will be considered as the concomitant medication started before the first dose of the investigational drug.

New concomitant medication

It is defined as a medication took on or after the first dose of the investigational drug to 90 days after the last dose of the investigational drug.

Previous and concomitant medications will be recorded on the corresponding CRF form. It will be coded according to WHO-Drug dictionary (version: WHODrug Global 2020 Mar or above), and summarized by ATC level 2, generic name and treatment group. For the same ATC classification and generic drug name, multiple drugs will be counted once only for a given subject.

6.5 Previous and Concomitant Non-drug Therapy

Previous non-drug therapy

It is defined as a non-drug therapy that has started and discontinued before the first dose of the investigational drug.

Concomitant non-drug therapy started before the first dose of the investigational drug

It is defined as the non-drug therapy that has started before the first dose of the investigational drug, and is continued after the first dose of the investigational drug. Non-drug therapy with unknown end date or therapy continuation will be considered as a concomitant non-drug therapy started before the first dose of the investigational drug.

New concomitant non-drug therapy

It is defined as a concomitant non-drug therapy used on or after the first dose of the investigational drug to 90 days after the last dose of the investigational drug.

Previous and concomitant non-drug therapy will be recorded on the following CRF form: previous and concomitant non-drug therapy, and previous surgical history-non-study disease. The verbatim terms will be coded using MedDRA dictionary (version 23.0 or above), and the number and percentage of subjects will be summarized by SOC and PT. For the same SOC and PT, multiple therapies of the same subject will be counted once only.

6.6 Subsequent New Anti-Tumor Therapy

Subsequent new anti-tumor therapy will be collected through the following CRF forms: subsequent anti-tumor therapy - surgery, subsequent anti-tumor therapy - chemotherapy, and subsequent anti-tumor therapy - radiotherapy. For the subsequent anti-tumor therapy, the following analysis will be carried out by treatment group:

- The number and percentage of the subjects receiving subsequent anti-tumor therapy at least once
- The number and percentage of the subjects receiving subsequent anti-tumor surgery. The subsequent anti-tumor surgery will be coded according to MedDRA dictionary (version 23.0 or above), and summarized by SOC, PT and treatment group (if applicable). For the same SOC and PT, multiple surgeries of the same subject will be counted once only.
- The number and percentage of the subjects receiving subsequent anti-tumor drug therapy. The subsequent anti-tumor drug therapy will be coded according to WHO-Drug dictionary (version: WHODrug Global 2020 Mar or above), and summarized by ATC level 2, generic name and treatment group. For the same ATC classification and generic drug name, multiple therapies will be counted once only for a given subject.
- The number and percentage of the subjects receiving subsequent anti-tumor radiotherapy, and the number and percentage of the subjects by radiotherapy site.
- The number and percentage of the subjects receiving subsequent anti-tumor immunotherapy. It will be coded according to WHO-Drug dictionary (version: WHODrug Global 2020 Mar or above), and summarized by ATC level 2, generic name and treatment group. For the same ATC classification and generic drug name, multiple drugs will be counted once only for a given subject.

Subsequent anti-tumor medications are recorded in the CRF page of "follow-up anti-tumor therapy — chemotherapy". Subsequent anti-tumor immunotherapy drugs will be identified and flagged by medical experts.

7 EFFICACY ANALYSIS

7.1 Primary Efficacy Analysis

The primary efficacy analysis will be conducted in intention-to-treat set (ITT). The PPS will also be used for the sensitivity analyses of primary efficacy endpoints (OS) and key secondary efficacy endpoints (PFS).

The efficacy data will also be provided in the individual subjects listing.

7.1.1 Definition of primary efficacy endpoints OS

Overall survival (OS) is defined as the time from randomization to death (for any cause). Subjects who are still alive at the time of analysis are censored at the known "last date of subject survival" or the "the analysis cut-off date", whichever is earlier. The "last date of subject survival" will be determined by searching all the confirmed dates of visits, evaluations, examinations, etc. collected in the database. See 4.2 for the imputation rules of partial or completely missing death date. Formula of OS: OS (month) = (date of death/censoring – date of randomization + 1)/30.4375.

7.1.2 Primary statistical analysis methods

The number and percentage of subjects with OS (death from any cause) events and censoring will be summarized by treatment group. The censoring reasons include survival at the analysis cutoff date, lost to follow-up, and withdrawal of ICF.

For the between-group OS comparison, the *p*-values are calculated using stratified log-rank test, with the stratification factors based on the actual data collected by eCRF and the data from the central laboratory.

Meanwhile, the between-group HR and corresponding 95% CI will be estimated using a stratified Cox proportional hazards model, with the stratification factors being the same as those in the log-rank test.

Kaplan-Meier method will also be used to analyze OS, and the survival curve will be estimated and plotted by treatment group. The median OS, the OS rates at 6, 12, 18 months and every 6 months thereafter and their 95% CIs will be calculated based on this curve.

The above mentioned OS primary analysis will be carried out in PD-L1 positive (CPS ≥ 5) population and ITT population, respectively. The stratification factors in ITT population are the same as those during randomization, including ECOG score, hepatic metastasis, and PD-L1 expression level. The stratification factors in PD-L1 positive population are ECOG score and hepatic metastasis.

7.1.3 Sensitivity analysis

The primary statistical analysis described in Section 7.1.2 will be repeated in the PPS set for sensitivity analysis carried out in PD-L1 positive (CPS ≥ 5) population and whole population.

An unstratified log-rank test will be performed for the between-group comparison, and an unstratified Cox proportional hazards model will be used to estimate the between-group HR and corresponding 95% CI.

7.1.4 Subgroup analysis

To evaluate the consistency of efficacy in various subgroups, subgroup analyses will be performed in the following factors (including but not limited to):

- Age (< 65 years or ≥ 65 years)
- Gender (male or female)
- Weight (< 60 kg or ≥ 60 kg)
- Randomization stratification factors: ECOG score (0 or 1), PD-L1 expression level (CPS < 10 or CPS ≥ 10), hepatic metastasis (yes or no)
- Tumor site (gastric, or gastroesophageal junction)
- Types of disease (locally progressive, or metastatic)
- Previous radical surgery (yes or no)
- PD-L1 expression level: CPS < 5 or CPS ≥ 5, CPS < 1 or CPS ≥ 1, TPS < 10% or TPS ≥ 10%, TPS < 5% or TPS ≥ 5%, TPS < 1% or TPS ≥ 1%

The number and percentage of OS events will be summarized by subgroup, the between-group HRs and corresponding 95% CI of efficacy endpoint OS in different subgroups will be estimated using an unstratified Cox proportional hazards model, and forest plots will be provided. The *p*-values of interaction effects between subgroup factors and treatment will also be provided. The following subgroup factors will not be included in the subgroup analysis of PD-L1 positive (CPS ≥ 5) population: PD-L1 expression level (CPS < 1 or CPS ≥ 1, CPS < 5, or CPS ≥ 5). Medical experts will determine whether a subject has received previous radical surgery by combining the information of previous surgery history and previous chemotherapy history.

7.2 Secondary Efficacy Analysis

7.2.1 Secondary efficacy analysis

The secondary efficacy endpoints will be statistically tested at a test level $\alpha = 0.05$ (two-sided), only when the primary efficacy endpoint is statistically significant.

Unless otherwise specified, the secondary efficacy analysis will be conducted in intention-to-treat set (ITT), and PPS will also be used for sensitivity analyses of PFS. In particular, the analyses of ORR, DCR and DoR will be carried out in the modified intention to treat (mITT) set.

7.2.2 Progression-free survival (PFS)

7.2.2.1 Definition of PFS

Progression-free survival (PFS) is defined as the time from randomization to the first recorded disease progression (imaging) or death for any reason, whichever occurs first. The subjects experiencing no progressive disease or death on or before the cutoff date of analysis will be censored (events and censoring rules are defined in Table 3).

Table 3. Events and Censoring rules for primary and sensitivity analysis of PFS

Scenarios	Primary analysis	Sensitivity analysis 1	Sensitivity analysis 2
Without imaging assessment or death record post baseline	Censored on the date of randomization	Censored on the date of randomization	Censored on the date of randomization
No PD or death, without initialing new anti-tumor therapy	Censored at the last imaging assessment	Censored at the last imaging assessment	If the subjects continue to receive randomized study treatment or complete 24 months of study treatment, or the sponsor terminates the study, the PFS data will be censored at the last complete imaging assessment. Otherwise they will be progressed at the investigational drug discontinuation.
Start of new anti-tumor therapy	Censored at the last imaging assessment before initialing the new anti-tumor therapy	Censored at the last imaging assessment before initialing the new anti-tumor therapy	If gastric cancer resection occurs after the first dose of the investigational drug and before treatment discontinuation, the PFS data will be censored at gastric cancer resection, otherwise they will be progressed at new anti-tumor treatment other than gastric cancer resection.
Death or progressive disease occurring after 1 missing complete imaging assessment	Progression at recorded date of death or PD	Progression at recorded date of death or PD	Progression at recorded date of death or PD

Scenarios	Primary analysis	Sensitivity analysis 1	Sensitivity analysis 2
Death or progressive disease occurring after ≥ 2 consecutive missing complete imaging assessments*	Censored at the last complete imaging assessment prior to consecutive missing	Progression at recorded date of death or PD	Progression at recorded date of death or PD
* According to the protocol, tumor imaging will be performed once every 6 weeks (± 7 days) for 48 weeks, and then once every 12 weeks (± 7 days). The interval between two consecutive tumor imaging assessments is defined as 91 days (13 weeks) before 48 weeks and 175 days (25 weeks) after 48 weeks. The definition of new anti-tumor treatment will be labeled by medical experts.			

7.2.2.2 Primary analysis of PFS

The primary analysis of PFS is similar to the statistical analysis of OS in Section 7.1.2, including the number of events (summarizing the number of disease progression and death events, respectively) and the number of censored subjects, stratified log-rank test, stratified Cox model to estimates the between-group HR and the 95% confidence intervals, survival curves for progression-free survival, median PFS estimates, and the PFS rate estimates at 6, 12, 18 months, and every 6 months thereafter and their 95% confidence intervals.

7.2.2.3 Sensitivity analysis of PFS

Sensitivity analysis will be conducted for different PFS censoring rules in Table 3.

The same PFS analysis described in 7.2.1 will be performed in the PPS. An unstratified log-rank test will be performed for the between-group comparison, and an unstratified Cox proportional hazards model will be used to estimate the between-group HR and 95% CI.

7.2.2.4 Subgroup analysis of PFS

It is the same as the subgroup analysis of primary efficacy endpoint.

7.2.3 Objective response rate (ORR)

Objective response rate (ORR) is defined as the proportion of subjects in the analysis set whose best overall response is CR or PR from the start of study treatment until the initiation of the new anti-tumor treatment during the study according to RECIST 1.1 criteria.

The ORRs and 95% confidence intervals will be calculated for each treatment group using normal approximation. The number and percentage of the best overall response will also be summarized by treatment group. The order of best overall response is CR, PR, stable disease (SD), progressive disease (PD), not evaluable (NE), and not evaluated. If the best overall response is SD, the SD shall be assessed at least 5 weeks after randomization, otherwise it shall be considered as NE.

The ORR and 95% confidence interval of the between-group difference and associated p-value will be calculated based on the stratified Miettinen-Nurminen methods where the stratification factors will be identical to those in the OS primary analysis. Subjects without any post baseline assessment will be considered as non-responders.

7.2.4 Disease control rate (DCR)

Disease control rate (DCR) is defined as the proportion of subjects in the analysis set whose best overall response is CR, PR and SD from the start of study treatment until the initiation of the new anti-tumor treatment.

The statistical analysis procedure for DCR is the same as that for ORR described in Section 7.2.3.

7.2.5 Duration of response (DoR)

Duration of response (DoR) is defined as the time from the first CR or PR to PD or death for any cause among the subjects with objective response (CR or PR) according to RECIST V1.1. The censoring rules for DoR are the same as those for PFS. See Table 3 primary analysis part in Section 7.2.2.1.

The number of cases (including the number of events and censoring), the maximum and minimum DoR (censoring or events), the median DoR, the probability of DoR at 3, 6, 9, 12 months and every 6 months thereafter and their 95% CIs (using the Kaplan-Meier method) will be summarized by treatment group, and the survival curves will be plotted.

The efficacy data by individual subjects will also be provided in the listing.

8 SAFETY ANALYSIS

All safety analyses will be performed in the safety set. The safety analysis is mainly based on the data during the treatment period, that is, from the first dose of the investigational drug until 90 days after the last dose of the investigational drug. Safety parameters include AEs, laboratory tests, vital signs, ECG, immunogenicity, etc.

8.1 Investigational Drug Exposure

The following information will be calculated for each investigational drug (IBI308/placebo, Capecitabine, and Oxaliplatin);

Exposure days

For (IBI308/placebo and Oxaliplatin), the exposure days will be calculated as: date of the last dose – date of the first dose + 21.

For Capecitabine, exposure days = date of the last dose – date of the first dose + 7.

Treatment duration (weeks)

The exposure duration (weeks) will be calculated as: exposure days/7, with one decimal place.

Treatment cycles

The number of doses (cycles) will be counted as the number of dosing cycles of the subject, and dose interruption is not considered.

Cumulative drug exposure

The cumulative drug exposure will be calculated as the sum of the actual dose in each cycle.

IBI308/ placebo exposure is calculated in mg, Capecitabine and Oxaliplatin exposure is calculated in mg/m^2.

Dose intensity per cycle

Dose intensity per cycle = cumulative drug exposure / number of cycles

i.e.

IBI308/placebo dose intensity (mg/cycle) = cumulative exposure of IBI308/placebo (mg) / number of IBI308/placebo cycles;

Capecitabine dose intensity (mg/m^2/cycle) = cumulative Capecitabine exposure (mg/m^2) / number of Capecitabine cycles;

Oxaliplatin dose intensity (mg/m^2/cycle) = cumulative Oxaliplatin exposure (mg/m^2) / number of Oxaliplatin cycles;

Relative dose intensity

Relative dose intensity (%) = 100 × (dose intensity per cycle / dose per cycle planned in the protocol) %

The doses planned in the protocol are shown in Table 4 (dose and administration), in which IBI308/placebo dose is calculated based on the baseline weight.

Table 4. Dose and administration

Investigational drug	Dose	Dosing frequency	Route of administration	Treatment Cycle	Usage
Sintilimab	Weight < 60 kg: 3 mg/kg	Q3W	Intravenous infusion	21 days as a cycle Administered on Day 1	Test group
	Weight ≥ 60 kg: 200 mg	Q3W	Intravenous infusion	21 days as a cycle Administered on Day 1	Test group
Placebo	Weight < 60 kg: 3 mg/kg	Q3W	Intravenous infusion	21 days as a cycle Administered on Day 1	Control group
	Weight ≥ 60 kg: 200 mg	Q3W	Intravenous infusion	21 days as a cycle Administered on Day 1	Control group
Oxaliplatin	130 mg/m^2	Q3W	Intravenous infusion	21 days as a cycle, administered on day 1, for 6 consecutive cycles	Test/Control group
Capecitabine	1000 mg/m^2	Q3W	Orally administered, 30 minutes after a meal	21 days as a cycle, for 14 consecutive days	Test/Control group

The exposure duration (weeks), number of doses (cycles) and the proportion of subjects who complete different cycles of Sintilimab, Capecitabine and Oxaliplatin in the study period will be summarized by treatment group. The above investigational drug exposure will also be summarized by subgroup (baseline weight ≥ 60 kg or < 60kg). The cumulative drug exposure, dose intensity per cycle, relative dose intensity and the proportion of subjects with relative dose intensity < 60%, 60–80%, 80–100%, 100–110%, and > 110% in the treatment period will also be summarized.

Compliance

Compliance = 100 × actual total dose/planned total dose, where the planned total dose of IBI308/placebo and Oxaliplatin is collected on the CRF, and the planned total dose of Capecitabine is derived from the administration date collected on the CRF. The calculation formula is: (end date of drug administration - start date of drug administration + 1) * planned total daily dose. The total daily dose is calculated based on BSA according to the following table:

Body surface area (BSA)	Total daily dose
[1.0,1.25)	2000
[1.25,1.50)	2500
[1.50,1.75)	3000
[1.75,2.0)	3500
2.0~	4000

Meanwhile, medication adherence is categorized by < 80%, 80–120%, and > 120%.

The reasons for the interruption of IBI308/placebo and Oxaliplatin infusion and the missed doses or adjusted doses of Capecitabine will be present in the listing.

8.2 Adverse Events

8.2.1 Definition

Treatment-emergent adverse events (TEAEs) are defined as any AEs that are occurred or worsen from the first dose of the investigational drug until 90 days after the last dose.

8.2.2 Planned analysis

All adverse events (AEs) will be coded according to MedDRA dictionary (version 23.0 or above), and graded according to CTCAE V5.0.

An overall summary table of AEs will be provided, in which the number and percentage of subjects with at least one AE in the following categories will be summarized by treatment group:

- All AEs
- All TEAEs
- Investigational drug-related TEAEs
- IBI308 or placebo-related TEAEs
- Capecitabine-related TEAEs
- Oxaliplatin-related TEAEs
- CTCAE Grade ≥ 3 TEAEs
- CTCAE Grade ≥ 3 Investigational drug-related TEAEs
- CTCAE Grade ≥ 3 IBI308 or placebo-related TEAEs
- CTCAE Grade ≥ 3 Capecitabine-related TEAEs
- CTCAE Grade ≥ 3 Oxaliplatin-related TEAEs
- SAEs
- Serious TEAEs
- Investigational drug-related serious TEAEs
- IBI308 or placebo-related serious TEAEs
- Capecitabine-related serious TEAEs

- Oxaliplatin-related serious TEAEs
- Immune-related TEAEs
- CTCAE Grade ≥ 3 Immune-related TEAEs
- Infusion reaction TEAEs
- Investigational drug-related infusion reaction TEAEs
- IBI308 or placebo-related infusion reaction TEAEs
- Oxaliplatin-related infusion reaction TEAEs
- TEAEs leading to death
- Investigational drug-related TEAEs leading to death
- IBI308 or placebo-related TEAEs leading to death
- Capecitabine-related TEAEs leading to death
- Oxaliplatin-related TEAEs leading to death
- TEAEs leading to discontinuation of any investigational drug
- TEAEs leading to discontinuation of IBI308 or placebo
- TEAEs leading to discontinuation of Capecitabine
- TEAEs leading to discontinuation of Oxaliplatin
- TEAEs leading to chemotherapy dose reduction
- TEAEs leading to Capecitabine dose reduction
- TEAEs leading to Oxaliplatin dose reduction
- TEAEs leading to interruption of any investigational drug
- TEAEs leading to interruption of IBI308 or placebo
- TEAEs leading to interruption of Capecitabine
- TEAEs leading to interruption of Oxaliplatin

The above AEs will also be summarized by baseline weight subgroup: < 60 kg or ≥ 60 kg.

The number and percentage of subjects with at least one AE and/or CTCAE Grade ≥ 3 AE in the following categories will also be summarized by SOC, PT and treatment group.

- TEAEs
- Investigational drug-related TEAEs
- IBI308 or placebo-related TEAEs
- Capecitabine-related TEAEs
- Oxaliplatin-related TEAEs
- Immune-related TEAEs
- Infusion reaction TEAEs
- Investigational drug-related infusion reaction TEAEs
- IBI308 or placebo-related infusion reaction TEAEs
- Oxaliplatin-related infusion reaction TEAEs
- Serious TEAEs
- Investigational drug-related serious TEAEs
- IBI308 or placebo-related serious TEAEs
- Capecitabine-related serious TEAEs
- Oxaliplatin-related serious TEAEs
- TEAEs leading to death
- Investigational drug-related TEAEs leading to death
- IBI308 or placebo-related TEAEs leading to death
- Capecitabine-related TEAEs leading to death
- Oxaliplatin-related TEAEs leading to death
- TEAEs leading to discontinuation of any investigational drug
- TEAEs leading to discontinuation of IBI308 or placebo
- TEAEs leading to discontinuation of Capecitabine
- TEAEs leading to discontinuation of Oxaliplatin
- TEAEs leading to chemotherapy dose reduction
- TEAEs leading to Capecitabine dose reduction

- TEAEs leading to Oxaliplatin dose reduction
- TEAEs leading to interruption of any investigational drug
- TEAEs leading to interruption of IBI308 or placebo
- TEAEs leading to interruption of Capecitabine
- TEAEs leading to interruption of Oxaliplatin
- Common TEAEs with an incidence ≥ 10%
- Common TRAEs with an incidence ≥ 5%

The number and percentage of subjects with TEAEs, investigational drug-related TEAEs, IBI308 or placebo-related TEAEs will be summarized by PT by treatment group. TEAEs with the difference of incidence rate ≥ 2% between two treatment group will be summarized by PT by treatment group.

Immune-related AEs and infusion reaction AEs will be summarized according to the investigator's judgment.

If multiple AEs of a same category occurred in a same subject, only the one with the highest CTCAE grade and the greatest relationship with the investigational drug is counted under this category. AEs with missing relationship with investigational drugs are considered as investigational drug-related.

All AEs, immune-related AEs, infusion reactions, SAE, AEs leading to drug interruption or discontinuation, CTCAE Grade ≥ 3 AE, and AEs leading to death will be listed, including start date, end date, study days, CTCAE grade, relationship with drugs, action taken, and outcome of AEs.

The death information of the subjects during the trial will be collected on the "Death" form in the CRF. The number and percentage of the subjects who died and the reason for death will be summarized.

8.3 Laboratory Tests

The summary of laboratory data will include hematology, blood chemistry, urinalysis, and other laboratory parameters (including coagulation function). Laboratory parameters are presented as below in Table 5.

Table 5. Laboratory parameters

Hematology	RBC, HGB, WBC, PLT, LYM, ANC
Blood chemistry	TBIL, ALT, AST, γ-GT, ALP, ALB, TP, LDH, UREA, Cr, Na, K, Cl, Mg, Ca, P, amylase, CK, CK-MB, and FBG
Urinalysis	pH, UWBC, UPRO, URBC, and UGLU
Coagulation function	PT、INR
Thyroid function	FT3、FT4、TSH
Virus serological test	HBsAg, HBsAb, HBcAb, HBeAg, HBeAb, HBV DNA, HCV antibody, treponema pallidum antibody, HCV RNA, and HIV antibody

ALB = albumin; ALP = alkaline phosphatase; ALT = alanine aminotransferase; ANC = absolute neutrophil count; AST = aspartate aminotransferase; Cr = serum creatinine; FBG = fasting blood glucose; FT3 = free triiodothyronine; FT4 = free thyroxine; γ-GT = γ-glutamyltransferase; UREA = urea; HBcAb = hepatitis B core antibody; HBeAb = hepatitis B e antibody; HBeAg = hepatitis B e antigen; HBsAb = hepatitis B surface antibody; HBsAg = hepatitis B surface antigen; HBV = hepatitis B virus; HCV = hepatitis C virus; HGB = hemoglobin; HIV = human immunodeficiency virus; INR = international normalized ratio; LDH = lactate dehydrogenase; LYM = lymphocyte count; PH = pH; PLT = platelet; PT = partial thromboplastin time; TBIL = total bilirubin; TP = total protein; TSH = thyroid stimulating hormone; RBC = red blood cell; UGLU = urine glucose; UPRO = urine protein; URBC = urine red blood cells; UWBC = urine white blood cells; WBC = white blood cell; Na = serum sodium ion; K = serum potassium ion; CL = serum chloride ion; Mg = serum magnesium ion; CA = serum calcium ion; P = serum phosphorus ion; CK = blood creatine kinase; CK-MB = blood creatine kinase isoenzyme.

Descriptive statistics of measurements of hematology, blood chemistry, coagulation function, thyroid function, and the changes from baseline will be summarized by treatment group and visit for continuous variables. For the category variable, a shift table will be used to summarize the normal and abnormal changes before and after treatment.

The severity of toxicities will be graded as per NCI Common Terminology Criteria for Adverse Events (CTCAE) version 5.0, if applicable. The baseline results and the worst results during the treatment period will be presented in a shift table.

For urinalysis, a shift table will be used to summarize the changes in categories with normal and abnormal at the scheduled post-baseline visit relative to the baseline.

The laboratory test results will be provided in the listing.

8.4 12-Lead ECG

The subjects are required to rest in supine position for at least 5 min prior to 12-lead ECG. The 12-lead ECG measurement schedules are shown in the schedule of visits in the protocol. The results of 12-lead ECG include the following parameters: QT interval, QRS, RR, and heart rate. QTcF is calculated using the following formula: QTcF = QT (msec)/(RR (sec)$^{1/3}$).

Descriptive statistics will be performed by treatment and visit for observed value and change from baseline for continuous parameters. For category parameters, a shift table will be used to summarize the normal and abnormal changes before and after treatment. The potentially clinical significant abnormalities during the treatment period will be summarized according to the classification shown in Table 6.

Table 6. Potential clinically significant abnormalities of ECG

ECG parameters	Category
QT	> 450–480 ms, > 480–500 ms, > 500 ms
Change in QT from baseline	> 30–60 ms, > 60 ms
QTcF	> 450–480 ms, > 480–500 ms, > 500 ms
Change in QTcF from baseline	> 30–60 ms, > 60 ms
QRS	≥ 120 ms
Heart rate	≤ 50 bpm and a decrease of ≥ 20 bpm from baseline ≥ 120 bpm and an increase of ≥ 20 bpm from baseline

8.5 Vital Signs, Physical Examination, and Other Safety Examinations

8.5.1 Vital signs

Vital signs include body temperature, pulse, respiratory rate, blood pressure, and body weight. The blood pressure and pulse in the supine position should be measured after subjects rest for at least 5 min. Vital signs should be measured before each administration. The vital signs measurement schedules are shown in the schedule of visits in the protocol. If multiple measurements are available at a same measurement time point, the mean of the multiple measurements will be used for analysis.

For the results of vital signs at each visit, body temperature, blood pressure (systolic and diastolic pressure), respiratory rate, pulse, and body weight will be summarized by treatment group using descriptive statistics, and change from baseline will be summarized too.

The potential clinically significant abnormalities of vital signs will be determined according to the threshold values given in Table 7. The number and percentage of subjects with at least one potential clinically significant abnormality during the treatment will be summarized.

Table 7. Potential clinically significant abnormalities of vital signs

Vital signs	Threshold values	Reference
Diastolic blood pressure	≤ 45 mmHg with a decrease of ≥ 10 mmHg from baseline ≥ 110 mmHg with an increase of ≥ 10 mmHg from baseline	All positions except standing
Systolic blood pressure	≤ 95 mmHg with a decrease of ≥ 20 mmHg from baseline ≥ 160 mmHg with an increase of ≥ 20 mmHg from baseline	All positions except standing
Heart rate	≤ 50 bpm with a decrease of ≥ 20 bpm from baseline ≥ 120 bpm with an increase of ≥ 20 bpm from baseline	All positions except standing
Body weight	An increase of ≥ 5% from baseline A decrease of ≥ 5% from baseline	FDA Feb 2007

8.5.2 Physical examination

A comprehensive physical examination includes: evaluations of general conditions, respiratory tract, cardiovascular system, abdomen, skin, head and neck (including ears, eyes, nose, and throat), lymph nodes, thyroid, musculoskeletal system (including spine and limbs), genitalia/anus (if required), and nervous system. A comprehensive physical examination is required at the screening visit, but not required in the subsequent treatment period. A physical examination will be carried out prior to drug administration at Day 1 of each cycle. For physical examination, changes in physical examination findings will be summarized by visit and treatment group. A detailed listing of physical examination results will be provided.

8.5.3 ECOG PS score

The ECOG performance status will be evaluated at screening, prior to drug administration at Day 1 of each cycle, the end-of-treatment visit, and the safety follow-up visit. The shift table of ECOG PS score will be provided before and after treatment.

8.5.4 Other safety examinations

Detailed listing will be provided for the results of other examinations.

8.6 Immunogenicity Analysis

Immunogenicity samples will be collected within 1 hour prior to Sintilimab/placebo infusion at Cycle 1/2/4/8/12/16, then every 8 cycles (Cycle 24, 32, etc. and so on) thereafter, and the safety follow-up visit. If an infusion-related reaction occurs during Sintilimab/placebo administration, blood samples should be taken near the start and end of this event and 30 days after the reaction, for comparative analysis of immunogenicity.

The subjects who have developed ADAs and NAbs during the study will be analyzed by baseline status (positive, negative, missing value, and total).

- The number and percentage of the subjects with ADA/NAb positive at baseline;

- The number and percentage of the subjects with treatment-emergent ADA/NAb positive; treatment-emergent positive is defined as negative at baseline and at least once ADA/NAb positive after baseline (treatment-induced positive), or positive at baseline and post-baseline ADA titer with 4-fold or more of baseline (treatment-boosted positive).
- The number and percentage of the subjects with transient ADA positive. Transient positive is defined as treatment-induced positive but not persistent positive.
- The number and percentage of subjects with persistent ADA positive. Persistent positive is a subset of treatment-induced positive, which is defined as at least 2 post-baseline positive, and the interval between the first and last positive samples is greater than or equal to 16 weeks; or only 1 post-baseline positive which is the last test result in the treatment period, or 1 post-baseline ADA positive at a time point less than 16 weeks before the last sample.
- The number and percentage of subjects at the earliest visit with positive results after baseline will be summarized.
- Post-baseline visits: provide the number and percentage of the subjects with ADA/Nab results at each visit after dosing;

The immunogenicity test results will be listed.

Depending on the ADA incidence rate, the impact of immunogenicity on safety and efficacy might be further analyzed, and the antibody titers might be summarized using descriptive statistics.

9 EXPLORATORY STATISTICAL ANALYSIS

9.1 Analysis of Predictive Biomarkers in Tumor Tissue and Peripheral Blood

When data are available, the proportion of subjects with different expression or distribution levels and the corresponding OS, PFS, ORR, DCR, and DoR might be summarized based on the overall baseline PD-L1 expression of the subjects.

9.2 Quality of Life Analysis

The changes in quality of life will be analyzed by treatment group in the safety set.

In this study, the quality-of-life evaluation will be performed on the day of the first dose, at each imaging evaluation visit, and the safety follow-up visit using the EQ-5D-5L, EORTC QLQ-C30, and EORTC QLQ-STO22 scales.

9.2.1 European quality of life - 5 dimensions - 5 levels (EQ-5D-5L)

European quality of life - 5 dimensions - 5 levels (EQ-5D-5L) is a multi-dimensional scale that measures health-related quality-of-life, which consists of 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and a visual analogue scale. Each dimension consists of 5 levels: no problems, slight problems, moderate problems, severe problems, and extreme problems. In addition, the index value is converted using the Japanese utility value conversion table based on the results of 5 dimensions.

For the 5 levels of the 5 dimensions, the number and percentage of subjects at each level will be provided by treatment group and visit.

For the visual scale and index value, the scores and changes from baseline will be summarized by treatment group and visit, and the curve of mean change from the baseline over time will be plotted.

9.2.2 European organization for research and treatment of cancer quality of life questionnaire (EORTC QLQ-C30)

QLQ-C30 scale includes 5 functional dimensions, 9 symptom dimensions or questions, and 1 global health status dimension. The formula of the raw score is shown in Table 8 below.

Table 8. Algorithm of Calculating Raw Score for QLQ-C30 scale

Domain (dimension)	Code	Property	Number of items	Score range (R)	Raw Score (RS)
Physical functioning	PF	Functional	5	3	$(Q_1+Q_2+Q_3+Q_4+Q_5)/5$
Role functioning	RF	Functional	2	3	$(Q_6+Q_7)/2$
Emotional functioning	EF	Functional	4	3	$(Q_{21}+Q_{22}+Q_{23}+Q_{24})/4$
Cognitive functioning	CF	Functional	2	3	$(Q_{20}+Q_{25})/2$
Social functioning	SF	Functional	2	3	$(Q_{26}+Q_{27})/2$
Global health status	QL		2	6	$(Q_{29}+Q_{30})/2$
Fatigue	FA	Symptom	3	3	$(Q_{10}+Q_{12}+Q_{18})/3$
Nausea and vomiting	NV	Symptom	2	3	$(Q_{14}+Q_{15})/2$
Pain	PA	Symptom	2	3	$(Q_9+Q_{19})/2$
Dyspnoea	DY	Symptom	1	3	Q_8
Insomnia	SL	Symptom	1	3	Q_{11}
Appetite loss	AP	Symptom	1	3	Q_{13}
Constipation	CO	Symptom	1	3	Q_{16}
Diarrhea	DI	Symptom	1	3	Q_{17}

Domain (dimension)	Code	Property	Number of items	Score range (R)	Raw Score (RS)
Financial difficulties	FI	Symptom	1	3	Q_{28}

The score in the last column of the table is the raw score derived from corresponding questionnaire. Missing values of questionnaire are not imputed, and not included in the numerators and denominators of the raw scores. The analysis variable is defined as the linear transformation of the original score, and the transformation rules are as follows:

For the functional dimension, the standard score is defined as $\left(1-\frac{\text{raw scores}-1}{R}\right)\times 100$;

For the symptom dimension and global health status, the standard score is defined as $\frac{\text{raw scores}-1}{R}\times 100$.

The value of standard score will range from 0 to 100. If more than half of the items are missing, the score of the dimension will be considered missing; the score of the dimension will be calculated only when the proportion of missing items is less than 50%.

The derived standard score and changes from the baseline will be summarized by treatment group and visit. The curve of mean change in global health status over time from the baseline will be plotted by treatment group and visit.

9.2.3 Quality-of-life questionnaire in patients with gastric cancer (EORTC QLQ-STO22)

The QLQ-STO22 comprises 22 items which are divided into 9 symptom dimensions: dysphagia, pain, reflux, eating restriction, anxiety, dry mouth, tasting, body image, and hair loss. The formula of the raw score is shown in Table 9 below.

Table 9. Algorithm of Calculating Raw Score for QLQ-STO22 scale

Domain (dimension)	Code	Property	Number of items	Score range (R)	Raw score (RS)
Dysphagia	STODYS	Symptom	3	3	$(Q_{31}+Q_{32}+Q_{33})/3$
Pain	STOPAIN	Symptom	4	3	$(Q_{34}+Q_{35}+Q_{36}+Q_{37})/4$
Reflux	STORFX	Symptom	3	3	$(Q_{38}+Q_{39}+Q_{40})/3$
Eating restriction	STOEAT	Symptom	4	3	$(Q_{41}+Q_{42}+Q_{43}+Q_{46})/4$
Anxiety	STOANX	Symptom	3	3	$(Q_{47}+Q_{48}+Q_{50})/3$
Dry mouth	STODM	Symptom	1	3	Q_{44}
Tasting	STOTA	Symptom	1	3	Q_{45}
Body image	STOBI	Symptom	1	3	Q_{49}
Hair loss	STOHL	Symptom	2	3	$(Q_{51}+Q_{52})/2$

The standard score derivations are the same as those for QLQ-C30 symptom dimensions.

The standard score and changes from baseline will be summarized by treatment group and visit, and the curve of mean change over time from the baseline will be plotted.

10 REFERENCES

1. Biostatistics guidelines for drug clinical trials, Center for Drug Evaluation, NMPA, 2016.
2. ICH E9 statistical principles for clinical trial. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), September 1998, ICH.
3. Common Terminology Criteria for Adverse Events (CTCAE) version 5.0. US Dept of Health and Human Services, National Institutes of Health, National Cancer Institute. 27November, 2017.
4. AS Pickard, MP Neary, D Cella: Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer Health Qual Life Outcomes 5: 70,2007
5. Maringwa JT, Quinten C, King M, Ringash J, Osoba D, Coens C, et al. Minimal important differences for interpreting health-related quality of life scores from the EORTC QLQ-C30 in lung cancer patients participating in randomized controlled trials. Support Care Cancer 2011;19:1753-60